DOCUMENT RESUME

ED 470 108 IR 021 546

AUTHOR Nelms, Keith R.

TITLE The Impact of Hypermedia Instructional Materials on Study Self-Regulation in College Students.

PUB DATE 2001-11-00

NOTE 7p.; In: Annual Proceedings of Selected Research and Development [and] Practice Papers Presented at the National Convention of the Association for Educational Communications and Technology (24th, Atlanta, GA, November 8-12, 2001). Volumes 1-2; see IR 021 504.

PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)

EDRS PRICE EDRS Price MF01/PC01 Plus Postage.

DESCRIPTORS *Computer Assisted Instruction; Computer Uses in Education; Educational Technology; Higher Education; *Hypermedia; *Instructional Material Evaluation; *Instructional Materials; Learning Processes; Nonprint Media; Printed Materials; Technology Integration

ABSTRACT

The metacognition "calibration of comprehension" research paradigm is used to investigate the question of whether the introduction of hypertext and hypermedia into college instruction impacts students' ability to regulate their own learning processes. Presentation technology (paper or computer) and content structure (linear or nonlinear) were independent variables in this 2x2 factorial design quantitative study. Instructional materials were differentially formatted to create the four experimental treatments: NP Environment, or nonlinear text in paper form (printed nonlinear Web site); LP Environment, or linear text in paper form (book); NC Environment, or nonlinear text on computer (hypermedia); and LC Environment, or linear text on computer. Subjects, undergraduate students at a small private college, were randomly assigned to the four treatment groups. Each treatment group had 17 subjects. After studying the treatment instructional materials, subjects predicted test performance on each of eight topics. Upon completion of an objective posttest, a comprehension calibration coefficient (the dependent variable) was calculated for each subject by correlating the eight performance predictions with the actual test scores on the eight topics using the Pearson product-moment correlation. Although statistically significant calibration was detected, analysis of variance found no statistically significant treatment or interaction effects. Data analysis and interviews suggest that subjects did not study at the same level of effort or use the same study strategies as used in real-world academic test preparation. It is not certain whether the findings of the experiment can be applied to actual college coursework environments. However, additional statistical analysis suggests that learning with unfamiliar media may impact calibration of comprehension for some students and that further investigation is needed. (Contains 22 references.) (Author/AEF)

Reproductions supplied by EDRS are the best that can be made
from the original document.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

P Harris

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

1

The Impact of Hypermedia Instructional Materials on Study Self-Regulation in College Students

Keith R. Nelms
Piedmont College

ED 470 108

Abstract

Does the introduction of hypertext and hypermedia into college instruction impact students' ability to regulate their own learning processes? The metacognition "calibration of comprehension" research paradigm is used to investigate this question. Interviews with experimental subjects provide additional insights into the study process.

Introduction

To be academically successful, college students must effectively allocate study effort among multiple courses based on the requirements of each course. An important study self-regulation skill is the ability to answer the pragmatic question "Do I know this subject matter well enough to take the test?" College students have developed this skill -- to varying degrees -- through years of studying paper-based textbooks. In today's college environment, Web and CD-ROM instructional materials require students to study materials displayed on a computer screen and organized in a nonlinear structure. Do these nonlinear hypertext and hypermedia instructional materials impact students' ability to accurately assess their own test readiness and thus to effectively regulate their study processes? If so, then the promotion of hypermedia instructional materials in the college environment may create unintended stumbling blocks to academic success.

Literature Review

The research literature to date does not address this issue. Although numerous studies have examined the use of hypertext and hypermedia as instructional media, there is no body of literature addressing impacts of hypermedia on self-regulation of the learning process. Some studies do, however, suggest there may be cause for concern.

Conklin (1987) noted two fundamental issues with hypermedia: disorientation and cognitive overhead. Hypermedia reading requires much greater mental effort in managing the reading process (compared to the simple page-turning of print environments); this mental activity can divert mental resources from the intellectual activity of reading and learning (Dede & Palumbo, 1991). Hypermedia users may lack the navigational skills needed to be successful in hypertext -based learning (Lawless & Kulikowich, 1996; Schroeder & Grabowski, 1995). Domain knowledge of the individual reader is a key determinant of a reader's ability to successful learning in the hypertext environment (Bernhuizen et al., 1994; Lawless & Kulikowich, 1996). Reading text from the computer screen generally requires more time (Belmore, 1985; Gould, Alfaro, Finn, Haupt, & Minuto, 1987; Grice, Ridgeway, & See, 1991; Kearsley, 1988). Some experiments have found poorer comprehension with computer-based text (Belmore, 1985; Feldmann & Fish, 1988; Fish & Feldmann, 1987; Reinking & Schreiner, 1985).

Although silent on the question of hypermedia impacts, reading and cognition researchers have investigated learning self-regulation. During the past 15 years, a number of research studies have explored learning self-regulation using a research paradigm known as "calibration of comprehension" (Lin & Zabrucky, 1998). In the current research, the calibration of comprehension paradigm has been adapted to investigate study self-regulation in a hypermedia learning environment.

In a typical calibration experiment, subjects read expository text and then are asked to predict their performance on a simple objective test over the materials read. Actual test performance is compared to self-assessed predicted performance using a correlation coefficient. Subjects able to accurately predict their performance are considered "highly calibrated" regardless of their performance on the test. Likewise, subjects who do not predict their performance accurately are considered "poorly calibrated."

Calibration of comprehension research has shown correlation between predicted performance and actual performance ranging from virtually zero to greater than $\underline{r} = .60$ (Lin & Zabrucky, 1998). Calibration research has also identified several ways in which research designs can maximize the probability of detecting calibration if it is indeed taking place. These guidelines were followed in the current research project:

1. Since posttest performance predictions may be influenced by subjects' prior knowledge of the topic (Glenberg & Epstein, 1987) and by subjects' interest in the topic (Glenberg et al., 1982; Lin et al., 1997), some method of assessing these subject attributes may be useful in the study.
2. Text should be of moderate difficulty for the research subject population (Weaver & Bryant, 1995; Weaver et al., 1995).
3. Posttests should have more than one question per text segment. Weaver (1990) found four questions per text to produce significantly more accurate indications of calibration than a single question per text.
4. Since subjects are better able to inventory their understanding and retention of facts than they are their ability to recognize logical inferences, posttest questions should deal with the recognition of facts and ideas (Glenberg et al., 1987; Pressley et al., 1987).

IR021546

BEST COPY AVAILABLE

Experiment

Presentation technology (paper or computer) and content structure (linear or nonlinear) were independent variables in this 2x2 factorial design quantitative study. As illustrated in Figure 1, instructional materials were differentially formatted to create the four experimental treatments. Content was identical in each treatment and consisted of eight topics. The instructional materials (and corresponding test questions) were adapted from a well-established curriculum and test bank.

		Presentation Technology	
		Paper (P)	**Computer (C)**
Content Structure	**Nonlinear (N)**	**NP Environment** Nonlinear text in paper form (printed nonlinear WWW site).	**NC Environment** Nonlinear text on computer (hypermedia).
	Linear (L)	**LP Environment** Linear text in paper form (book).	**LC Environment** Linear text on computer.

Figure 1. Experimental treatments created by varying presentation technology and content structure.

Experimental subjects (undergraduate students at a small private college) were randomly assigned to the four treatment groups. Each treatment group had 17 subjects. After studying the treatment instructional materials, subjects predicted test performance on each of the eight topics. Upon completion of an objective posttest, a comprehension calibration coefficient (the dependent variable) was calculated for each subject by correlating the eight performance predictions with the actual test scores on the eight topics using the Pearson product-moment correlation.

As noted above, the research literature suggests that subject interest in the study topic, subject expertise in the topic, and subject motivation to perform in the posttest may all be covariates with the calibration coefficient. Length of study time might also reasonably be a covariate. Based on researcher-recorded reading times and subject self-reported measures of interest, expertise, and motivation, no significant covariance was found.

To assess the impact of presentation technology and content structure on subject calibration of comprehension, data generated in the experiments were subjected to hypothesis testing:

1. $H_0$1: There was no significant difference between the calibration coefficients for the computer technology treatment and the paper technology treatment. (Rejection of this hypothesis would mean the presentation technology influences calibration.)
2. $H_0$2: There was no significant difference between the calibration coefficients for linear structure treatment and the nonlinear structure treatment. (Rejection of this hypothesis would mean the linear/nonlinear structure of the instructional materials influences calibration.)
3. $H_0$3: There was no significant interaction effect between the technology and structure treatments as measured by the calibration coefficient. (If hypothesis $H_0$3 were to be rejected, then three more hypotheses would be tested.)
4. $H_0$3a: There was no significant difference between the calibration coefficients for the linear paper treatment and the nonlinear computer treatment. (This hypothesis compared a typical book format to nonlinear computer hypermedia.)
5. $H_0$3b: There was no significant difference between the calibration coefficients for the nonlinear computer treatment and the nonlinear paper treatment. (This hypothesis compared calibration when reading from a website to calibration when reading from a printed copy of the website.)
6. $H_0$3c: There was no significant difference between the calibration coefficients for the nonlinear computer treatment and the linear computer treatment. (This hypothesis compared two different design approaches for hypermedia.)

Experimental Results

The dependent variable, referred to as the "calibration coefficient", was a Pearson Product-Moment Correlation calculated between subjects' self-predicted performance on eight fallacy topics with their actual posttest scores on those topics. The mean value of the calibration coefficient is 0.09 which is significantly greater than zero, $t(67) = 1.95$, $p<0.05$. The median value of the calibration coefficient variable is 0.15. Figure 2 displays the distribution of the calibration coefficient for the experimental subjects.

BEST COPY AVAILABLE

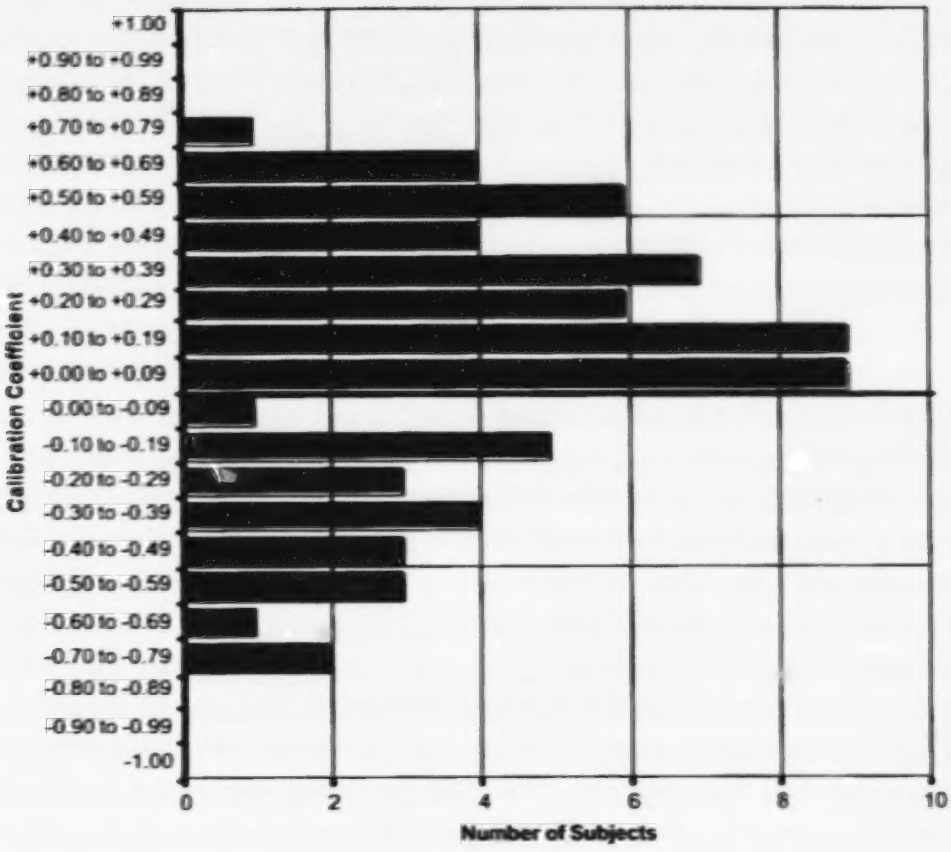

Figure 2. Distribution of dependent variable.

An analysis of variance was conducted to determine if treatments or treatment interactions affected subjects' ability to predict test performance and thus regulate their study processes. As noted in Table 1, treatment effects were not statistically significant nor were there statistically significant interaction effects. Thus, the first three hypotheses are not rejected and the second three hypotheses are not tested.

Table 1. ANOVA with calibration coefficient as dependent variable

Source	SS	df	MS	F
Structure	.176	1	.176	1.251
Technology	.020	1	.020	.146
2-way Interaction	.050	1	.050	.354
Residual	9.005	64	.141	

Post hoc power analysis with $p<0.05$ indicates a power level less than 0.20

Several aspects of the experimental data warranted further investigation:

1. The mean of the calibration coefficients, while statistically greater than zero, was disappointingly low. The 0.9 mean is reminiscent of the earliest calibration research and was expected to be higher since the research meticulously followed guidelines developed in the calibration research literature.
2. The distribution of the calibration coefficient shown in Figure 2 is disturbing. It has the general appearance of a random normal distribution centered on zero and shows a large number of "negatively calibrated" subjects. Negative calibration coefficients imply that subjects consistently score poorly on topics they think they know well and vice versa. Negative correlation is not addressed in the calibration of comprehension literature and has no obvious ties to real world learning

experiences. The negative correlation found in the calibration coefficient distribution suggests a major random process underlying the calibration coefficient.

3. The test questions (along with the instructional texts) have an extensive track record at the host institution. The grade distributions typical for the posttest questions were well known. Posttest scores were lower and more disperse than expected.
4. Reliability analysis and item analysis indicated problems with the test questions. Since extensive past use of the instructional materials and test questions suggest these materials are effective and appropriate together, then low reliability and item analysis scores indicate some other problem in the study methodology.

Taken together, these observations pointed to extensive random guessing by subjects during the posttest phase of the experiment. The randomizing influence of posttest guessing would produce the symptoms noted above.

Interviews

To obtain further insights into the experimental results, two subjects from each of the four treatment groups were selected for interviews. For each treatment group, one subject had a high posttest score and a relatively large positive calibration coefficient. The second subject for each treatment group had a high posttest score and a relatively large negative calibration coefficient. (Only subjects with high posttest scores were selected to make sure the interviewed subjects were actively engaged in the experiment's learning task.)

A thirteen-question telephone interview was conducted with each of the eight identified research subjects. The questions addressed several issues arising from analysis of the experimental data:

1. Effort. Seven of the eight interviewed subjects admitted they would have put much more effort into studying the instructional materials if a course grade had been at stake. Six of the eight subjects estimated guessing at 20%-40% of the posttest questions; one estimated guessing at 60% of the questions.
2. Stopping criteria. Interviewed subjects were ask to describe their study "stopping criteria" – how they decided when to stop studying – for both real world study tasks and for the experimental task. Only three of the eight interviewed subjects used the same criteria in the experiment as they typically used when studying for college courses.
3. Anti-calibration. None of the eight interviewed subjects reported study self-regulation difficulties that resemble anti-calibration.

Based on these interviews, it is clear that subjects were not motivated to study the experimental materials in the same way or to the same extent they study actual academic materials. It can not be assumed then, that the subjects engaged their normal calibration skills either. The extensive guessing during the experimental posttest introduced a large-scale random influence to the experimental data. Thus, it is not clear to what extent the experimental data accurately reflects student calibration in real world settings.

Other Findings

The interviews revealed four types of study stopping criteria:

1. Process criteria. Several subjects described study "rituals" involving reading/re-reading practices, note-taking, or other study strategies. Process-oriented students tended to stop studying once the study process was complete.
2. Feel good criteria. Some subjects reported studying until they felt they knew the materials.
3. Feel bad criteria. Other subjects reported studying until they felt like they weren't getting anything more out of the studying.
4. Time criteria. Some subjects reported studying until they ran out of time and the test was given.

Subjects reported a variety of ergonomic and interface concerns. Eye strain and navigational confusion were mentioned by subjects reading from the computer screen. One subject noted her "body got bored" reading the hypertext – she found the instructional material interesting, but became physically restless having to sit in one position in front of the computer screen. Reading from a book would allow her to change physical positions and read for longer periods of time. Subjects noted they were unable to highlight text on the computer like they did in textbooks. One subject who visualized the book during testing found he could not use the same memory technique when reading from the computer. Two subjects suggested hypermedia would be an excellent tool for reviewing materials already read.

Figure 3 presents one of the most interesting observations from the experimental data. The scattergram shows individual subjects plotted by posttest score performance (x-axis) and predicted performance (y-axis). Regression lines for each of the four treatment groups are shown, along with the R^2 for each regression line.

In examining Figure 3, it is obvious that the LP treatment (the treatment representing traditional studying from a textbook) is noticeably different from the other treatments. The regression line and R^2 for the LP treatment suggests that students who scored better on the posttest were also better able to predict their performance than students who did not mastered the material. This is intuitively appealing, regardless of whether this indicates that better calibrated students learn more or that students who have learned more are better able to assess their knowledge. Note, however, that the other treatments show no such relationship. Since the discussion above has already established that the artificial experimental setting heavily influenced the research data, one cannot presume Figure 3 necessarily represents impacts of the treatments in real world study situations. However, these results do suggest further study is warranted.

BEST COPY AVAILABLE

Summary

The present study is a pioneering foray into an unexplored issue--the calibration of comprehension in hypermedia environments. In particular, this study was designed to determine if hypermedia instructional materials would impact college students' ability to assess their own readiness for testing.

Two basic characteristics of hypermedia--content structure and presentation technology--were used as variables to define a 2 x 2 experimental design. The content structure variable could be either linear (L) or nonlinear (N); the presentation technology variable could be either paper-based (P) or computer-based (C). The variables define four treatment categories: linear paper-based materials (LP), linear computer-based materials (LC), nonlinear paper-based materials (NP), and nonlinear computer-based materials (NC).

After reading the instructional materials, but before seeing the posttest questions, subjects were asked to predict their test performance on questions for each of the eight topics contained in the instructional materials. The posttest generated eight different scores--one for each of eight topics contained in the experimental materials. For each subject, a Pearson Product-Moment Correlation (Pearson r) was calculated by pairing the eight posttest scores with the eight score predictions. These Pearson r values--called calibration coefficients--served as the dependent variable in the experiment.

Although statistically significant calibration was detected, analysis of variance found no statistically significant treatment or interaction effects. Data analysis and interviews suggest that subjects did not study at the same level of effort or use the same study strategies as used in real-world academic test preparation. This lack of preparation resulted in greater levels of guesswork during the posttest and thus increased randomness in the experimental data. Since students may not have used the same study strategies in the experiment as they use in actual academic coursework, it is not certain whether the findings of the experiment can be applied to actual college coursework environments. However, additional statistical analysis suggests that learning with unfamiliar media may impact calibration of comprehension for some students and that further investigation is needed. Future research seeking to address calibration in academic settings should be incorporated into academic coursework where possible.

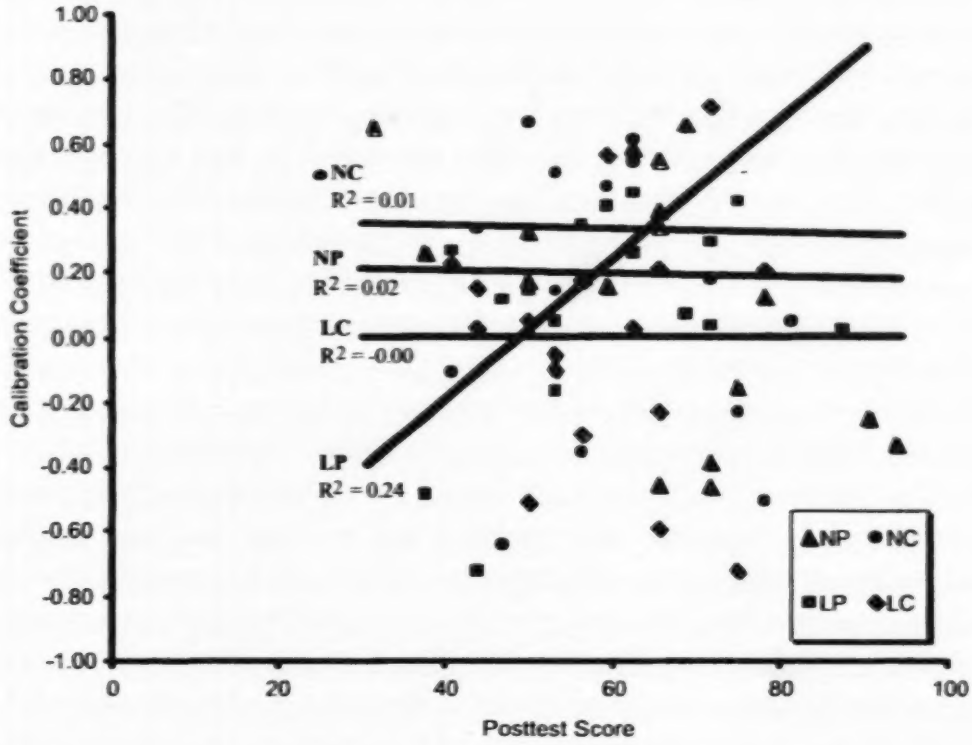

Figure 3. Scattergram of calibration coefficients and posttest scores with regression lines by treatment.

BEST COPY AVAILABLE

References

Bonhuizen, J., Stoutjesdijk, E., & van Putten, K. (1994). Studying textbooks: Effects of learning styles, study task, and instruction. Learning and Instruction, 4(2), 151-174.

Belmore, S. M. (1985). Reading computer-presented text. Bulletin of the Psychonomic Society, 23, 12-14.

Conklin, J. (1987). Hypertext: An introduction and survey. Computer, 20(9), 17-41.

Dede, C. J., & Palumbo, D. B. (1991). Implications of hypermedia for cognition and communication. Impact Assessment Bulletin, 9(1-2), 15-27.

Feldmann, S. C., & Fish, M. C. (1988). Reading comprehension of elementary, junior high and high school students on print vs. microcomputer-generated text. Journal of Educational Computing Research, 4, 159-166.

Fish, M. C., & Feldmann, S. C. (1987). A comparison of reading comprehension using print and microcomputer presentation. Journal of Computer-Based Instruction, 14, 57-61.

Glenberg, A. M., & Epstein, W. (1987). Inexpert calibration of comprehension. Memory & Cognition, 15(1), 84-93.

Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. Journal of Experimental Psychology: General, 116(2), 119-136.

Glenberg, A. M., Wilkinson, A., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. Memory & Cognition, 10, 597-602.

Gould, J. D., Alfaro, L., Finn, R., Haupt, B., & Minuto, A. (1987). Reading from CRT displays can be as fast as reading from paper. Human Factors, 29, 497-517.

Grice, R. A., Ridgeway, L. S., & See, E. J. (1991). Hypertext: Controlling the leaps and bounds. Technical Communications (First Quarter), 48-56.

Kearsley, G. (1988). Authoring considerations for hypertext. Educational Technology (November), 21-24.

Lawless, K. A., & Kulikowich, J. M. (1996). Understanding hypertext navigation through cluster analysis. Journal of Educational Computing Research, 14(4), 385-399.

Lin, L., & Zabrucky, K. M. (1998). Calibration of comprehension: Research and implications for education and instruction. Contemporary Educational Psychology, 23, 345-391.

Lin, L., Zabrucky, K., & Moore, D. (1997). The relations among interest, self-assessed comprehension, and comprehension performance in young adults. Reading Research and Instruction, 36(2), 127-139.

Maki, R. H., Foley, J. M., Kajer, W. K., Thompson, R. C., & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(4), 609-616.

Pressley, M., Snyder, B. L., Levin, J. R., Murray, H. G., & Ghatala, E. S. (1987). Perceived readiness for examination performance (PREP) produced by initial reading of text and text containing adjunct questions. Reading Research Quarterly, XXII(2), 219-236.

Reinking, D., & Schreiner, R. (1985). The effects of computer-mediated text on measures of reading comprehension and reading behavior. Reading Research Quarterly, 20, 536-552.

Schroeder, E. E., & Grabowski, B. L. (1995). Patterns of exploration and learning with hypermedia. Journal of Educational Computing Research, 13(5), 313-335.

Weaver, C. A., III. (1990). Constraining factors in calibration of comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(2), 214-222.

Weaver, C. A., III, & Bryant, D. S. (1995). Monitoring of comprehension: The role of text difficulty in metamemory for narrative and expository text. Memory & Cognition, 23(1), 12-22.

Weaver, C. A., III, Bryant, D. S., & Burns, K. D. (1995). Comprehension monitoring: Extension of the Kintsch and van Dijk model. In C. A. I. Weaver, S. Mannes, & C. R. Fletcher (Eds.), Discourse comprehension: Essays in honor of Walter Kintsch (pp. 177-193). Hillsdale, NJ: Erlbaum.

BEST COPY AVAILABLE

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

X	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").